


# Inferring the population properties of binary black holes from unresolved gravitational waves

Rory J. E. Smith<sup>1,2</sup> ,<sup>1,2</sup>★ Colm Talbot,<sup>1,2,3</sup> Francisco Hernandez Vivanco<sup>1,2</sup> and Eric Thrane<sup>1,2</sup>

<sup>1</sup>*School of Physics and Astronomy, Monash University, VIC 3800, Australia*

<sup>2</sup>*OzGrav: The ARC Centre of Excellence for Gravitational Wave Discovery, Clayton, VIC 3800, Australia*

<sup>3</sup>*LIGO, California Institute of Technology, Pasadena, CA 91125, USA*

Accepted 2020 June 2. Received 2020 May 25; in original form 2020 April 21

## ABSTRACT

The vast majority of compact binary mergers in the Universe produce gravitational waves that are too weak to yield unambiguous detections; they are unresolved. We present a method to infer the population properties of compact binaries – such as their merger rates, mass spectrum, and spin distribution – using both resolved and unresolved gravitational waves. By eliminating entirely the distinction between resolved and unresolved signals, we eliminate bias from selection effects. To demonstrate this method, we carry out a Monte Carlo study using an astrophysically motivated population of binary black holes. We show that some population properties of compact binaries are well constrained by unresolved signals after about one week of observation with Advanced LIGO at design sensitivity.

**Key words:** gravitational waves – black hole mergers.

## 1 INTRODUCTION

Every year, around  $2 \times 10^6$  binary neutron stars and  $1.5 \times 10^5$  binary black holes (BBHs) merge somewhere in the Universe, radiating gravitational waves (Abbott et al. 2018a). Only a small fraction of these signals are detected by observatories such as Advanced LIGO (aLIGO), Advanced Virgo, and KAGRA (Acernese et al. 2014; Aasi et al. 2015; Akutsu et al. 2019). The rest are too faint to be resolved. None the less, the ensemble of unresolved gravitational-wave signals forms an astrophysical background, which *can* be detected by advanced gravitational-wave detectors (Abbott et al. 2016b, 2018a; Smith & Thrane 2018; Hernandez-Vivanco et al. 2019). Here, we use the word ‘background’ to denote gravitational-wave signals that are not clearly detected and published in catalogues, for example Abbott et al. (2019). Since there are many connotations associated with the notion of a gravitational-wave background, it is worth pausing to make our meaning absolutely clear.

First, we note that this definition of ‘background’ is detector-dependent; as gravitational-wave detectors become more sensitive, a greater fraction of binary mergers will be clearly resolved, and so what we might refer to as background now will become foreground in the future. Second, we note that the gravitational-wave background from compact binaries is often thought of as a *foreground* when looking for primordial gravitational waves from the early Universe; see, for example Maggiore (2000). Indeed, one scientist’s foreground is another’s background; here we use the

word ‘foreground’ to refer to resolved binaries. Finally, there is a common notion that the gravitational-wave background consists of a plethora of unimaginably faint sources. In reality, it derives from a continuum of binaries, ranging from the nearly detectable to the clearly not detectable. Since there is no universally accepted definition of ‘detection,’ the boundary between the resolved catalog and the unresolved background is fuzzy.

However one may choose to delineate this boundary, the background encodes rich information about the mass and spin distributions of compact binaries. These distributions, in turn, provide insights into binary evolution (Stevenson, Ohme & Fairhurst 2015; Farr et al. 2017; Gerosa & Berti 2017; Stevenson, Berry & Mandel 2017; Talbot & Thrane 2017; Vitale et al. 2017; Lower et al. 2018; Wysocki, Lange & O’Shaughnessy 2019), star formation history, the fate of massive stars (Fishbach & Holz 2017; Abbott et al. 2018a; Talbot & Thrane 2018), the behaviour of matter at supranuclear densities (Abbott et al. 2018b), and the existence of primordial black holes (Raidal, Vaskonen & Veermäe 2017), amongst other things. Crucially, the foreground probes only the closest binaries. By analysing the foreground and background together, it is possible to probe the entire population of binary mergers.

Here, we use hierarchical inference<sup>1</sup> to extend the method outlined in Smith & Thrane (2018) in order to determine the ensemble properties of compact binaries. By eliminating the artificial distinction between foreground and background, we probe

\* E-mail: rory.smith@monash.edu

<sup>1</sup>For a review of hierarchical inference in gravitational-wave astronomy, see section V of Thrane & Talbot (2019).

**Table 1.** Hyper-parameters  $\Lambda_i$  of the BBH mass and spin population distributions.

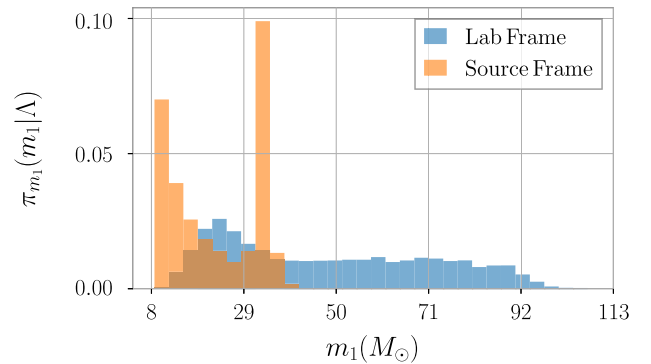
Hyper-parameter $\Lambda_i$	Description	Injection value
$\xi$	Astrophysical duty cycle	$6.67 \times 10^{-3}$
$m_{\min} (M_{\odot})$	Minimum black hole mass	$8.68 M_{\odot}$
$m_{\max} (M_{\odot})$	Maximum mass of black holes in the power-law component	$39.5 M_{\odot}$
$\mu_m (M_{\odot})$	Mean of the Gaussian component of the primary mass distribution	$33.4 M_{\odot}$
$\sigma_m$	Standard deviation of the Gaussian component of the primary mass distribution	$1.08 M_{\odot}$
$\lambda_m$	Fraction of black holes in the Gaussian component of the primary mass distribution	0.340
$\alpha_m$	Slope of the power-law component of the primary mass distribution	2.00
$\beta_m$	Slope of the mass ratio distribution	-0.198
$a_{\max}$	Maximum spin magnitude	1.00
$\alpha_a$	Spin-magnitude beta distribution slope parameter (rise)	1.50
$\beta_a$	Spin-magnitude beta distribution slope parameter (fall)	3.50
$\sigma_{\text{tilt}}$	Standard deviation of the spin-tilt angle distribution	1.00
$\xi_{\text{tilt}}$	Fraction of BBHs with Gaussian distributed spin tilts	0.50

greater distances than possible with resolved events alone, while eliminating bias from selection effects. We demonstrate that it is possible to make population inferences even when excluding statistically significant, ‘gold-plated’ detections. The key results are posterior probability distributions describing the shape of the BBH mass and spin distributions, derived using entirely unresolved events. We show that these posteriors are consistent with the true values used for the generation of the simulated data. We argue that this method is statistically optimal in the sense that it is not possible to obtain more narrow posteriors given a fixed data set.

This work builds on Gaebel et al. (2019), which describes how population studies can be extended to include subthreshold candidate events, some of which are bona fide gravitational-wave signals, even though any single candidate is probably a noise fluctuation. This is part of a broader trend in gravitational-wave astronomy. For example, the arguably marginal event GW170729 was included<sup>2</sup> in the first gravitational-wave transient catalog GWTC-1 (Abbott et al. 2018b) and the companion paper (Abbott et al. 2018a).

We highlight a few innovations unique to this work. First, we eliminate selection effects entirely by making no distinction between detected events and subthreshold events. Taking into account selection effects in population studies can be a somewhat subtle endeavour (Abbott et al. 2016a; Fishbach, Holz & Farr 2018; Mandel, Farr & Gair 2019; Thrane & Talbot 2019), involving challenging efficiency calculations (Ng et al. 2018; Tiwari, Fairhurst & Hannam 2018). These challenges are removed by eliminating the concept of a detection threshold. Second, by eliminating the minimum detection threshold entirely, we extend the range of the analysis to include events at large redshifts, well beyond what can be probed with unambiguous detections. This is an important first step toward studying the evolution of binary populations over cosmic time, though, more work is required to measure this redshift-dependence using hyper-parameters; see Fishbach et al. (2018). Third, while Gaebel et al. (2019) generate pseudo-posterior samples from a Fisher matrix approximation for the likelihood function, we calculate posterior samples using a full-fledged parameter estimation pipeline. By carrying out full parameter estimation (the main computational cost of the search), we show that our method is computationally feasible.

The remainder of this paper is organized as follows. In Section 2, we describe astrophysically motivated models of the BBH mass

**Figure 1.** Astrophysically motivated primary mass ( $m_1$ ) distribution in the source frame (orange) and lab frame (blue). The lab-frame mass distribution appears redshifted due to the expansion of the universe.

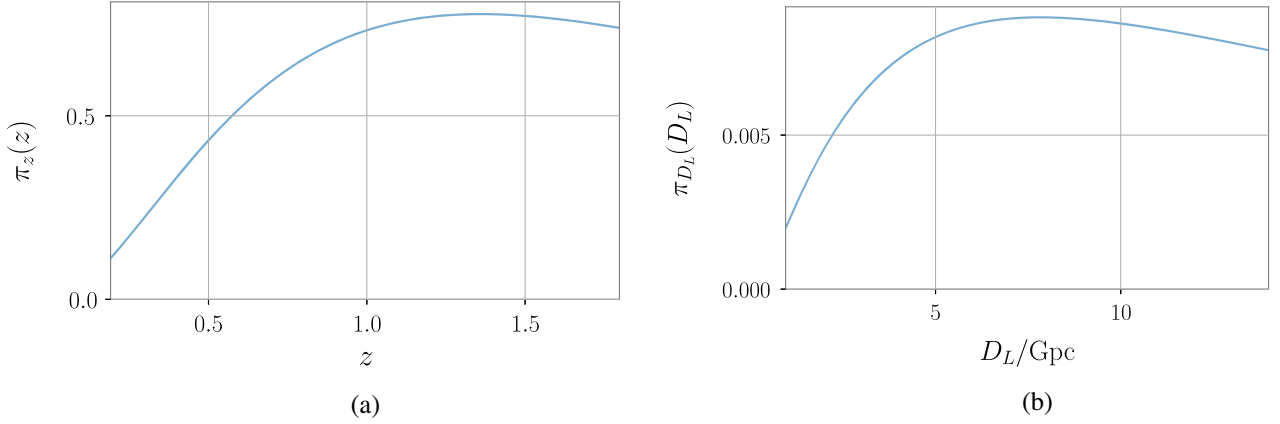
spectrum and spin distributions. In Section 3, we describe the method for population inference from a population of subthreshold signals. In Section 4, we present the results of our Monte Carlo study. Concluding remarks are provided in Section 6.

## 2 POPULATION MODEL

We parametrize the mass and spin distributions using one of the prescriptions from Abbott et al. (2018a). In this section, we briefly summarize our population model. The reader is referred to the appendix for more details. Our models take the form of conditional priors  $\pi_{\theta}(\theta|\Lambda)$  where  $\theta$  are BBH parameters and  $\Lambda$  are hyper-parameters governing the shape of the  $\theta$  distribution. A list of hyper-parameters, their meaning, and injection values used in this study is provided in Table 1.

We model the black hole mass spectrum following Talbot & Thrane (2018). The distribution is a mixture model of a truncated power law and a Gaussian. An example of the source-frame primary mass distribution is shown in orange in Fig. 1 and the lab-frame distribution (distorted but cosmological redshift) is shown in blue. We model the distribution of black hole spin magnitudes following Wysocki et al. (2019). The distribution is a beta distribution. We model the distribution of black hole spin orientations following Talbot & Thrane (2017). The distribution is a mixture model of an isotropic distribution and model with a preference for aligned spin. For this study, we choose a set of plausible population parameters based on Abbott et al. (2018a).

<sup>2</sup>The event GW170729 has an astrophysical probability ranging from  $p_{\text{astro}} = 48 - 98$  per cent.



**Figure 2.** Prior distributions on redshift (left) and luminosity distance (right).

We assume a fixed, known redshift distribution of (or equivalently, luminosity distance). We assume that sources are uniformly distributed in comoving volume to a maximum luminosity distance of  $d_L^{\max} \approx 5$  Gpc (redshift  $z = 0.8$ ). Throughout, we assume the standard  $\Lambda$ -cold dark matter ( $\Lambda$ CDM) cosmology ( $\Omega_\Lambda = 0.69$ ,  $\Omega_m = 0.31$ , and  $H_0 = 67.7 \text{ km Mpc}^{-1} \text{ s}^{-1}$ , Ade et al. 2016). While this distance distribution ignores effects arising from the time-dependent star formation rate, see Fishbach et al. (2018) and You et al. (2020), it is satisfactory for our present purposes. By probing redshifts up to  $z = 0.8$  (lookback time = 7 Gyr), it is in-principle possible to glean information about a time when the Universe was younger and the star formation rate was higher Madau & Dickinson (2014). In Figs 2(a) and (b), we show the explicit redshift and luminosity distributions implied by our uniform-in-comoving volume source distribution with standard  $\Lambda$ CDM cosmology.

The final ingredient required to characterize our population of BBHs is the duty cycle  $\xi$ , the fraction of segments containing a BBH signal. In the next section, we describe how the data are divided into 16 s segments. Current observations of BBH mergers suggest that two black holes merge somewhere in the Universe on average once every  $223^{+352}_{-115}$  s. Most of these mergers probably take place at redshifts of  $z < 2$  ( $d_L \lesssim 15$  Gpc). Beyond  $z = 2$ , it is believed that star formation rate decreases (Madau & Dickinson 2014). With fewer stars, there are fewer black holes, and therefore fewer mergers. Assuming an average time between BBH of 100 s out to  $d_L = 15$  Gpc, the duty cycle out to luminosity distances of 5 Gpc is approximately  $\xi = 6.67 \times 10^{-3}$ , and so we use this value for our injection study.

### 3 INFERENCES FROM THE GRAVITATIONAL-WAVE BACKGROUND

#### 3.1 Overview

This section describes the statistical formalism that allows us to calculate the hyper-posterior distribution  $p(\Lambda|\mathbf{d})$  for population parameters  $\Lambda$  described in Section 2 given some data set  $\mathbf{d}$ . We follow the method described in Smith & Thrane (2018). The calculation is divided into the following steps.

(i) We divide the data into 16 s segments. These segments are a convenient size so that any given segment is unlikely to contain more than one BBH signal. However, they are long enough that it is relatively unlikely for a BBH signal to fall on the boundary of two

segments; see Smith & Thrane (2018). Assuming an average merger rate of one per 223 s the probability of obtaining two merger signals in one 16 s segment is approximately  $(16 \text{ s}/223 \text{ s})^2 \approx 0.5$  per cent.

(ii) Run the nested sampling code DYNesty (Speagle 2020, implemented in the `bilby`, Ashton et al. 2018, Bayesian inference library) to generate posterior samples  $\{\theta_{k,i}\}$  describing the mass and spins of individual BBH events in each segment. Additionally, DYNesty estimates for each data segment, the noise evidence  $\mathcal{Z}(d|\mathcal{H}_N)$  – that there is no BBH present – and ‘the default signal evidence’  $\mathcal{Z}(d|\mathcal{H}_S)$  – that there is a BBH signal present given some default prior  $\pi(\theta)$ .

(iii) The posterior samples and evidences for each segments are used to define a ‘total likelihood’, defined in equation (1), which combines data from many segments. We discuss the hyper-likelihood in greater detail in the next subsection.

(iv) Having defined the hyper-likelihood, we use DYNesty to generate hyper-posterior samples  $\{\Lambda_i\}$ , which provide a representation of  $p(\Lambda|\mathbf{d})$ .

Steps (i) and (ii) are relatively straightforward. In the next subsection, we describe the hyper-likelihood used in steps (iii) and (iv).

#### 3.2 The hyper-likelihood

Following Smith & Thrane (2018), we employ a likelihood function to describe the probability of some large data set  $\mathbf{d}$  given a population of BBH described by hyper-parameters  $\xi$  (the fraction of data segments containing a signal) and  $\Lambda$ , which describes the shape of the BBH mass and spin distributions

$$\mathcal{L}^{\text{tot}}(\mathbf{d}|\Lambda, \xi) = \prod_i^n [\xi \mathcal{L}(d_i|\Lambda, \mathcal{H}_S) + (1 - \xi) \mathcal{Z}(d_i|\mathcal{H}_N)]. \quad (1)$$

There is a lot to explain in this equation and the rest of this subsection is devoted to this task. The  $\text{tot}$  superscript denotes that this is the likelihood for the entire data set  $\mathbf{d}$ . The expression includes a product over  $i$  data segments running from  $i = 1$  to  $n$ . The term  $\mathcal{L}(d_i|\Lambda, \mathcal{H}_S)$  is the single-segment Bayesian evidence for the data  $d_i$  in segment  $i$  given the signal hypothesis  $\mathcal{H}_S$  and hyper-parameters  $\Lambda$ . The term  $\mathcal{Z}(d_i|\mathcal{H}_N)$  is the single-segment noise evidence for the data  $d_i$  in segment  $i$ . The hyper-parameter  $\xi$  is often referred to as ‘duty cycle,’ and may be converted into a rate (Smith & Thrane 2018).

The single-segment noise evidence  $\mathcal{Z}(d_i|\mathcal{H}_N)$  is straightforwardly calculated for each segment using a Gaussian-noise

likelihood<sup>3</sup>

$$\mathcal{Z}(d_i|\mathcal{H}_N) = \exp\left(-\frac{1}{2}\langle d_i, d_i \rangle\right). \quad (2)$$

Here, we employ a noise-weighted inner product

$$\langle a, b \rangle \equiv 4\pi\Delta f \sum_k \frac{a^*(f_k)b(f_k)}{S_n(f_k)}, \quad (3)$$

where the sum is over frequency bins  $k$  with bin widths of  $\Delta f$  and  $S_n(f)$  is the strain noise power spectral density.

The single-segment signal likelihood  $\mathcal{L}(d_i|\Lambda, \mathcal{H}_S)$  is given by (equation 5) yielding:

$$\mathcal{L}(d_i|\Lambda, \mathcal{H}_S) \approx \frac{\mathcal{Z}(d_i|\mathcal{H}_S)}{n_s} \sum_{k=1}^{n_s} \frac{\pi(\theta_{k,i}|\Lambda)}{\pi(\theta_{k,i})}. \quad (4)$$

Here,  $\mathcal{Z}(d|\mathcal{H}_S)$  is the Bayesian evidence for a BBH signal in segment  $i$  calculated using some default prior for the BBH parameters  $\theta$  denoted  $\pi(\theta)$ . Assuming Gaussian noise, it is given by

$$\begin{aligned} \mathcal{Z}(d_i|\mathcal{H}_S) &\equiv \int d\theta_i \mathcal{L}(d_i|\theta, \mathcal{H}_S)\pi(\theta_i) \\ &= \int d\theta_i \exp\left(-\frac{1}{2}\langle d_i - h(\theta_i), d_i - h(\theta_i) \rangle\right) \\ &\quad \times \pi(\theta_i), \end{aligned} \quad (5)$$

where  $h(\theta)$  is the gravitational waveform, in this case, calculated IMRPhenomPv2 approximant (Hannam et al. 2014; Smith et al. 2016). The integral in equation (5) is calculated numerically using the Bayesian inference library, *bilby* (Ashton et al. 2018) implementation of DYNesty (Speagle 2020). In addition to calculating  $\mathcal{Z}(d_i|\mathcal{H}_S)$ , *bilby* outputs a list of  $n_s$  posterior samples  $\{\theta_{k,i}\}$ , which describe the posterior  $p(\theta_i|d_i)$  given the default prior. It is sometimes said that the ratio of priors  $\pi(\theta_{k,i}|\Lambda)/\pi(\theta_{k,i})$  in equation (4) serves to ‘reweight’ the posterior samples calculated using the default prior  $\pi(\theta)$  (Thrane & Talbot 2019).

### 3.3 The hyper-posterior

Using the hyper-likelihood defined in equation (1), it is straightforward to obtain the (hyper-) posterior for duty cycle and the other hyper-parameters  $\Lambda$

$$p(\Lambda, \xi|\mathbf{d}) = \frac{\mathcal{L}^{\text{tot}}(\mathbf{d}|\Lambda, \xi)\pi(\Lambda)\pi(\xi)}{\mathcal{Z}_\Lambda^{\text{pop}}}. \quad (6)$$

Here  $\pi(\Lambda)$  is the hyper-parameter prior, which we take to be uniform for each hyper-parameter. The distribution  $\pi(\xi)$  is the duty cycle prior. In a real analysis, one should choose a distribution, which uses a Poisson distribution to relate duty cycle to astrophysical rate; see Smith & Thrane (2018). However, for our present purposes, it is convenient to simply employ a uniform prior. The variable  $\mathcal{Z}_\Lambda^{\text{pop}}$  is the hyper-evidence. They hyper-evidence can be used to carry out model selection between different population models; see Talbot & Thrane (2018), Stevenson et al. (2015), Stevenson et al. (2017), Abbott et al. (2018a), Stevenson et al. (2017), Vitale et al. (2017), Talbot & Thrane (2017), Gerosa & Berti (2017), Farr et al. (2017), Wysocki et al. (2019), and Lower et al. (2018).

<sup>3</sup>We note that this is missing a normalization factor, however, as this only depends on the PSD and not on the template, we can freely factor this out of the both the signal and noise evidences.

**Table 2.** Priors on the 15 BBH signal parameters,  $\pi(\theta)$ . The priors are used in Stage 1 of the hierarchical population inference (Section 3). The parameters are the source-frame primary black hole mass,  $m_1$ ; mass ratio  $q$ ; comoving distance  $D_C$ ; time of coalescence  $t_c$ ; cosine of the orbital inclination  $\cos i$ ; phase at coalescence  $\phi_c$ ; polarization phase  $\psi$ ; cosine of the spin-tilt angles  $\cos t_1$  and  $\cos t_2$ ; the angle between the two spin vectors  $\phi_{12}$ ; angle between the total and orbital angular momentum  $\phi_{JL}$ ; dimensionless spin magnitudes  $a_1$  and  $a_2$ ; right ascension  $\alpha$ ; and cosine of the declination  $\delta$ .

Parameter $\theta_i$	Prior $\pi(\theta_i)$
$m_1$	Uniform( $6 M_\odot, 50 M_\odot$ )
$q$	Uniform( $0.2, 1$ )
$D_C^3$	Uniform( $1 \text{ Gpc}^3, 5^3 \text{ Gpc}^3$ )
$t_c$	Uniform( $0\text{s}, 16\text{s}$ )
$\cos i$	Uniform( $-1, 1$ )
$\phi_c$	Uniform( $0, 2\pi$ )
$\psi$	Uniform( $0, \pi$ )
$\cos t_1$	Uniform( $-1, 1$ )
$\cos t_2$	Uniform( $-1, 1$ )
$\phi_{12}$	Uniform( $0, 2\pi$ )
$\phi_{JL}$	Uniform( $0, 2\pi$ )
$a_1$	Uniform( $0, 1$ )
$a_2$	Uniform( $0, 1$ )
$\alpha$	Uniform( $0, 2\pi$ )
$\cos \delta$	Uniform( $-1, 1$ )

## 4 RESULTS: DEMONSTRATION WITH SIMULATED DATA

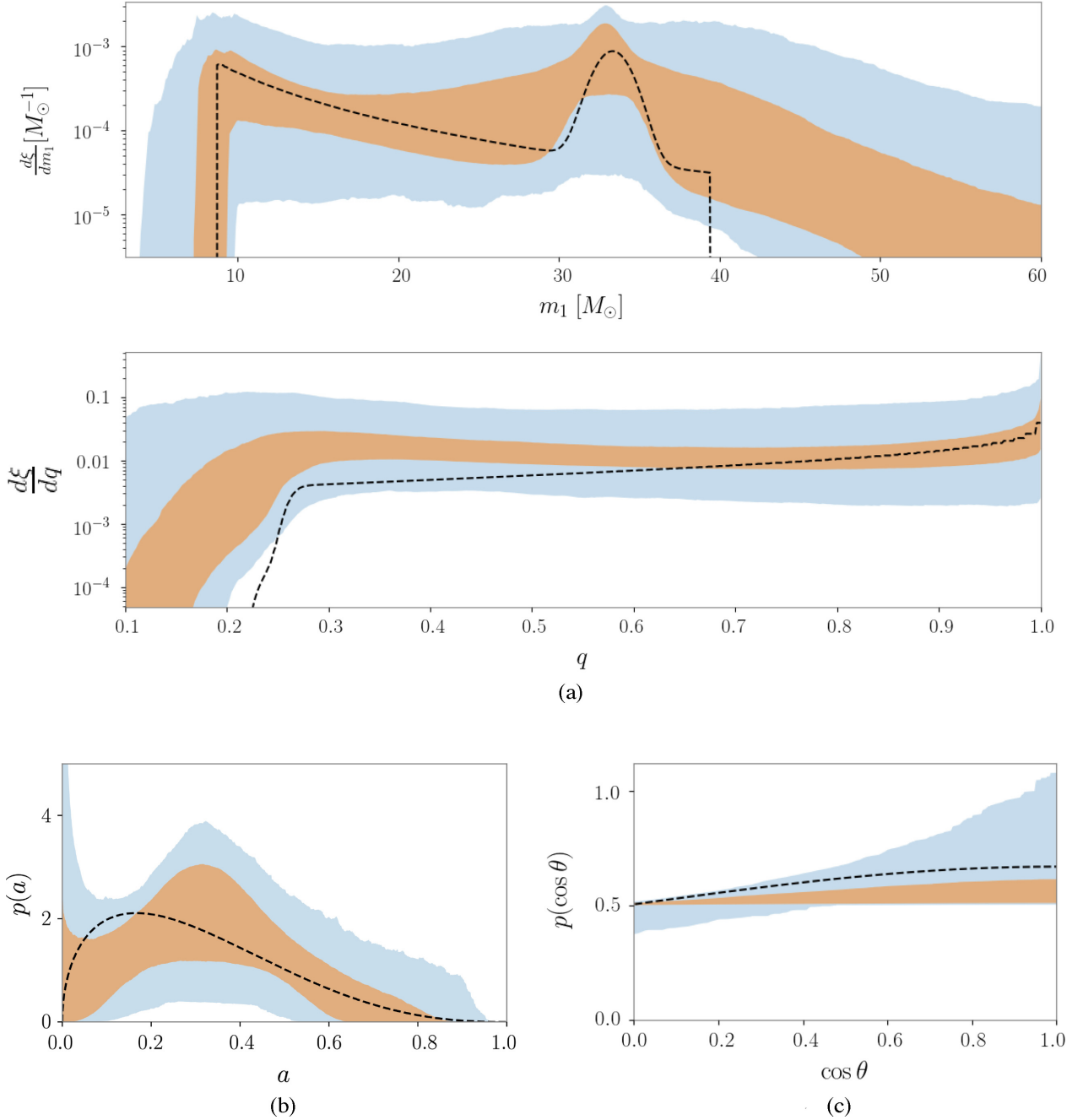
We analyse 5.5 d of simulated aLIGO design-sensitivity data (Aasi et al. 2015) containing an ensemble of 200 simulated BBH signals. We divide the data into  $3 \times 10^4$  16-s segments. This yields a duty cycle  $\xi = 200/30000 = 6.67 \times 10^{-3}$ . We derive the duty cycle by first assuming an average merger range of BBHs of 1 per 100 s. We then assume that the merger rate drops significantly beyond a redshift of  $z \sim 2$  so that their contribution can be effectively ignored. The fraction of all binaries contained in the volume with maximum redshift considered here,  $z = 0.8$ , is approximately 4 per cent. The average merger rate out to  $z = 0.8$  is then approximately one merger per 45 min. In 5.5 d, this yields 176 binary mergers, however we choose to round up to 200.

The masses and spins of the BBHs are drawn from the mass and spin distributions described in Section 2. The remaining ‘extrinsic’ parameters are drawn using standard distributions. All of the signals in our injection set are below the usual threshold for matched-filter network S/N (signal-to-noise ratio):  $\rho_{\text{network}}^{\text{th}} = 12$ . Based on results from Smith & Thrane (2018), we expect the BBH background to be detectable with approximately one day of aLIGO design sensitivity data.

We estimate the signal and noise evidence  $\mathcal{Z}_S, \mathcal{Z}_N$ , and obtain posterior samples for BBH source parameters for every data segment. The priors, summarized in Table 2, and are chosen to be relatively uninformative so we can recycle the posterior samples later. We then use the sets of evidence and posterior samples as input to equation (3.3) to compute the posterior for  $\Lambda$  – the population mass and spin distribution parameters – and  $\xi$ , the astrophysical duty cycle.

The computational cost of running full parameter estimation on  $3 \times 10^4$  16-s data segments is kept manageable by explicitly marginalizing over three parameters, which are difficult to sample: comoving distance, coalescence time, and coalescence phase; see, for example Thrane & Talbot (2019) for the details of these





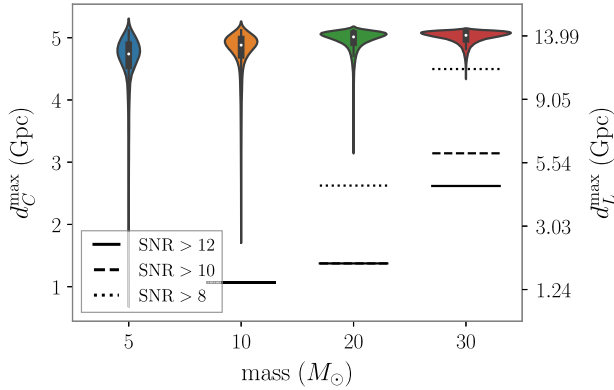
**Figure 3.** Posterior predictive distributions of BBH parameters. These results are obtained using five and a half days of simulated aLIGO data containing 395 BBHs signals. The dashed line is the true distribution, while the red contours represent the 50 per cent (light) and 90 per cent credible intervals on the inferred distributions. The parameters are: (top) primary black hole mass  $m_1$ , (centre) mass ratio  $q$ , (lower left) spin magnitude  $a$ , and (lower right) cosine spin tilt  $\cos \theta$ .

marginalization schemes. By marginalizing over these parameters, we significantly decrease the convergence time, and hence run time, of computing evidences and drawing posterior samples in step 1.

We find that the background is detectable within one week out to comoving distances of 5 Gpc, assuming masses and spins drawn from the distribution described in Section 2. The posterior distribution on  $\xi$  is consistent with the true value of  $\xi = 0.67$  per cent, and the log Bayes factor (equation 15 of Smith & Thrane 2018) overwhelmingly supports a detection of a population of compact binaries:  $\ln \text{BF} \approx 700$ , confirming the previous result from Smith & Thrane (2018) with a different, more realistic population of BBH.

We find that we can begin to constrain some of the mass and spin population parameters are using the the 200 unresolved mergers in our simulated data. In Fig. 3(a), we show posterior predictive distributions for different mass and spin parameters. The posterior predictive distributions reflect our updated prior based on information from our hyper-posteriors; see Thrane & Talbot (2019). The contours represent the  $1\sigma$  and  $2\sigma$  credible intervals.

In Fig. A1, we show posterior distributions for hyper-parameters associated with the duty cycle and mass parameters. In Fig. A2, we show posterior distributions for the parameters associated with the Gaussian component of the mass population model. In Fig. A3, we



**Figure 4.** Violin plots of the comoving and luminosity  $d_{\max}$  posterior obtained by running hierarchical inference with different mass distributions. Each distribution is set to be Gaussian with standard deviation  $\sigma = 0.3 M_{\odot}$  and means  $\mu = (5, 10, 20, 30) M_{\odot}$ . The horizontal solid, dashed, and dotted lines correspond to the most distant event observed with network S/N > 12, > 10, and > 8 for each mass, respectively. The posteriors peak at the true value  $d_{\max} = 5100$  Mpc (comoving distance) and the most distant events with S/N > 12 lie below the  $d_{\max}$  posteriors, suggesting that we obtain most of the information from subthreshold events.

show posterior distributions for hyper-parameters describing black hole spins.

## 5 HOW SENSITIVE ARE WE TO SUBTHRESHOLD EVENTS?

In this section, we investigate where the information for our analysis comes from. Is our resolving power coming primarily from binaries just below the detection threshold, or do we gain information from weaker events as well? To address this question, we carry out a follow-up study where we introduce a new hyper-parameter,  $d_{\max}$ , the maximum comoving distance for binary mergers. In our new population model, the rate of binary mergers drops to zero for distances greater than  $d_{\max}$ . The  $d_{\max}$  parameter is not physical, but it is useful for our present investigation: if the data disfavour some value of  $d_{\max}$  (less than the true value of  $d_{\max}$ ), then we are getting information from that distance. We set the true value of  $d_{\max} = 5100$  Gpc (comoving distance) and then use hierarchical inference to obtain a posterior for  $d_{\max}$ . We calculate the posterior on  $d_{\max}$  for different Gaussian mass distributions with standard deviation  $\sigma = 0.3 M_{\odot}$  and means  $\mu = (5, 10, 20, 30) M_{\odot}$ . The results are shown in Fig. 4.

The posterior on  $d_{\max}$  peaks at the true value of  $d_{\max} = 5100$  Mpc (comoving distance). The  $d_{\max}$  likelihood is clearly informative for distances greater than the distance of the furthest S/N > 12 event, which is marked by the horizontal solid black line in Fig. 4. This is true for all masses considered in our study. A similar conclusion is made for the most distant event with S/N > 10, marked by the horizontal dashed line and the most distant event with S/N > 8, marked the horizontal dotted line. (No events with S/N > 12 were used to obtain this hyper-posterior.) This plot is a good indication that we are indeed getting information from subthreshold events.

## 6 CONCLUSIONS

Our results demonstrate that the astrophysical gravitational-wave background can be used to constrain the population properties of BBHs together with ‘gold plated’ foreground signals. By ap-

plying hierarchical inference to all available data – irrespective of whether it contains a gravitational-wave signal or not – we eliminate selection bias. By carrying out population inferences with subthreshold events we help extend the reach of the current generation of observatories to greater distances. A crucial next step is the demonstration of the algorithm using real data. A mock data challenge is underway to show how the algorithm performs in realistic conditions. Another goal is to determine how much information can be inferred about the redshift dependence of BBH mass and spin properties.

It is also interesting to compare the results, for example from Fig. 3(a), with results from LIGO’s second observing run (Abbott et al. 2018a). In our analysis, we exclude 20 resolvable detections with  $\rho_{\text{network}} > 12$ , which is twice the number analysed in Abbott et al. (2018a). The simulated constraints shown here are qualitatively similar, though, likely not nearly as constraining as the actual results from Abbott et al. (2018a). In this example, it would seem that the inclusion of unresolved events is likely to yield a modest, though non-negligible improvement in population inferences, albeit with a population of significantly more distant sources.

## ACKNOWLEDGEMENTS

RS, CT, FHV, and ET are supported by the Australian Research Council (ARC) CE170100004. ET is supported by ARC FT150100281. We thank Stuart Anderson and the LIGO Data Grid for assistance with computing infrastructure, and Maya Fishbach, Thomas Callister, and Thomas Dent for helpful comments and suggestions. We acknowledge the OzStar cluster for providing graphical processor units to carry out some of our calculations.

## REFERENCES

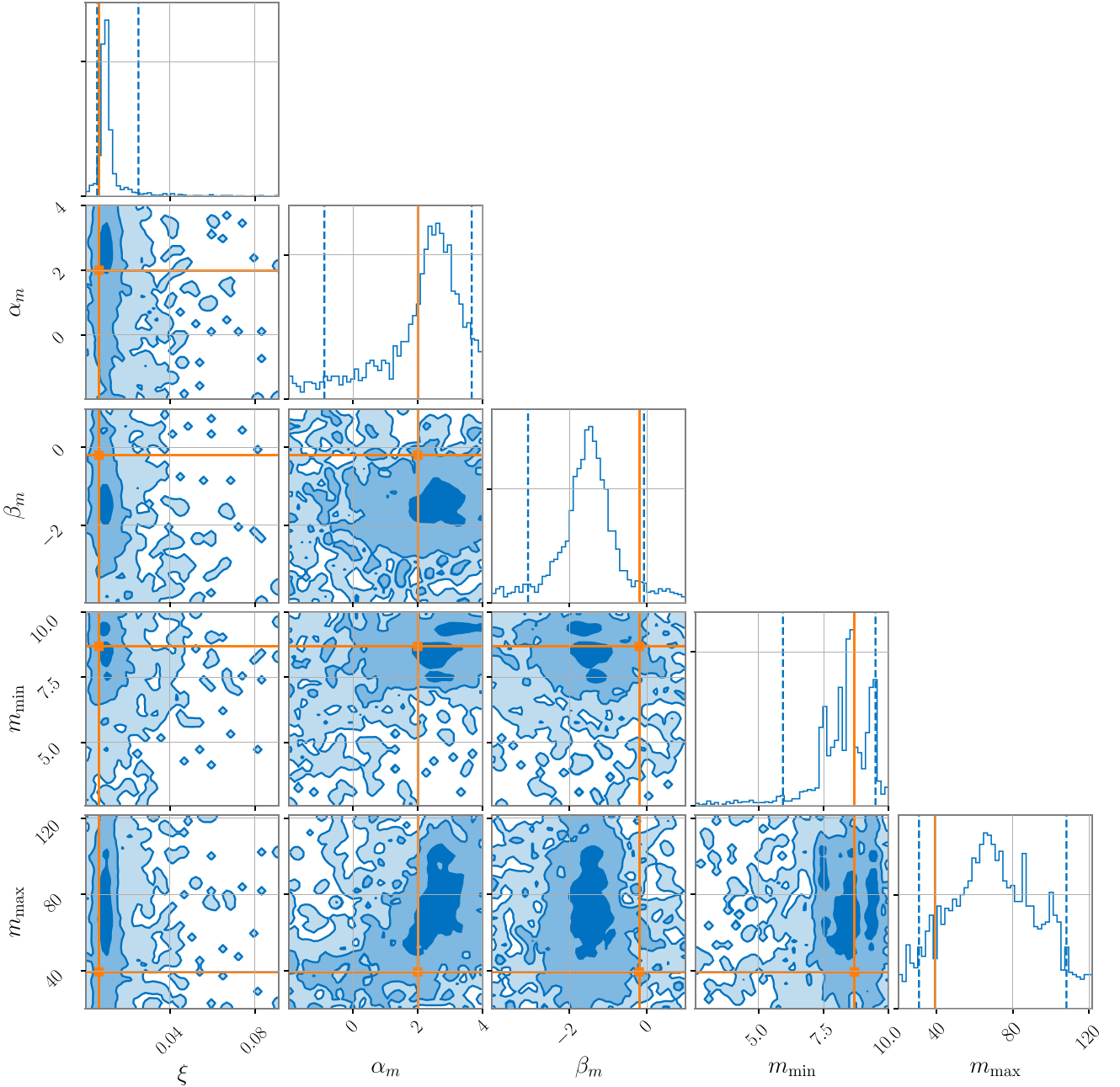
- Aasi J. et al., 2015, *Class. Quantum Gravity*, 32, 074001
- Abbott B. P. et al., 2016a, *Phys. Rev. X*, 6, 041015
- Abbott B. P. et al., 2016b, *Phys. Rev. Lett.*, 116, 131102
- Abbott B. P. et al., 2018a, *Phys. Rev. Lett.*, 120, 091101
- Abbott B. P. et al., 2018b, *Phys. Rev. Lett.*, 121, 161101
- Abbott B. P. et al., 2019, *Phys. Rev. X*, 9, 031040
- Acernese F. et al., 2014, *Class. Quantum Gravity*, 32, 024001
- Ade P. A. R. et al., 2016, *A&A*, 594, A13
- Akutsu T. et al., 2019, *Nat. Astron.*, 3, 35
- Ashton G. et al., 2018, preprint ([arXiv:1811.02042](https://arxiv.org/abs/1811.02042))
- Farr W. M., Stevenson S., Miller M. C., Mandel I., Farr B., Vecchio A., 2017, *Nature*, 548, 426
- Fishbach M., Holz D. E., 2017, *ApJ*, 851, L25
- Fishbach M., Holz D., Farr B., 2017, *ApJ*, 840, L24
- Fishbach M., Holz D. E., Farr W. M., 2018, *ApJ*, 863, L41
- Gaebel S. M., Veitch J., Dent T., Farr W. M., 2019, *MNRAS*, 484, 4008
- Gerosa D., Berti E., 2017, *Phys. Rev. D*, 95, 124046
- Hannam M., Schmidt P., Bohé A., Haegel L., Husa S., Ohme F., Pratten G., Pürrer M., 2014, *Phys. Rev. Lett.*, 113, 151101
- Hernandez-Vivanco F., Smith R. J. E., Thrane E., Lasky P. D., 2019, *Phys. Rev. D*, 100, 043023
- Lower M. E., Thrane E., Lasky P. D., Smith R., 2018, *Phys. Rev. D*, 98, 083028
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Maggiore M., 2000, *Phys. Rep.*, 331, 283
- Mandel I., Farr W. M., Gair J. R., 2019, *MNRAS*, 486, 1086
- Marchant P., Renzo M., Farmer R., Pappas K. M. W., Taam R. E., de Mink S. E., Kalogera V., 2019, *ApJ*, 882, 36
- Ng K. K. Y., Vitale S., Zimmerman A., Chatziioannou K., Gerosa D., Haster C.-J., 2018, *Phys. Rev. D*, 98, 083007

- Raidal M., Vaskonen V., Veermäe H., 2017, *J. Cosmol. Astropart. Phys.*, 2017, 037
- Smith R., Thrane E., 2018, *Phys. Rev. X*, 8, 021019
- Smith R., Field S. E., Blackburn K., Haster C.-J., Pürrer M., Raymond V., Schmidt P., 2016, *Phys. Rev. D*, 94, 044031
- Speagle J. S., 2020, *MNRAS*, 493, 3132
- Stevenson S., Ohme F., Fairhurst S., 2015, *ApJ*, 810, 58
- Stevenson S., Berry C., Mandel I., 2017, *MNRAS*, 471, 2801
- Talbot C., Thrane E., 2017, *Phys. Rev. D*, 96, 023012
- Talbot C., Thrane E., 2018, *ApJ*, 856, 173
- Thrane E., Talbot C., 2019, *PASA*, 36, e010
- Tiwari V., Fairhurst S., Hannam M., 2018, *ApJ*, 868, 140

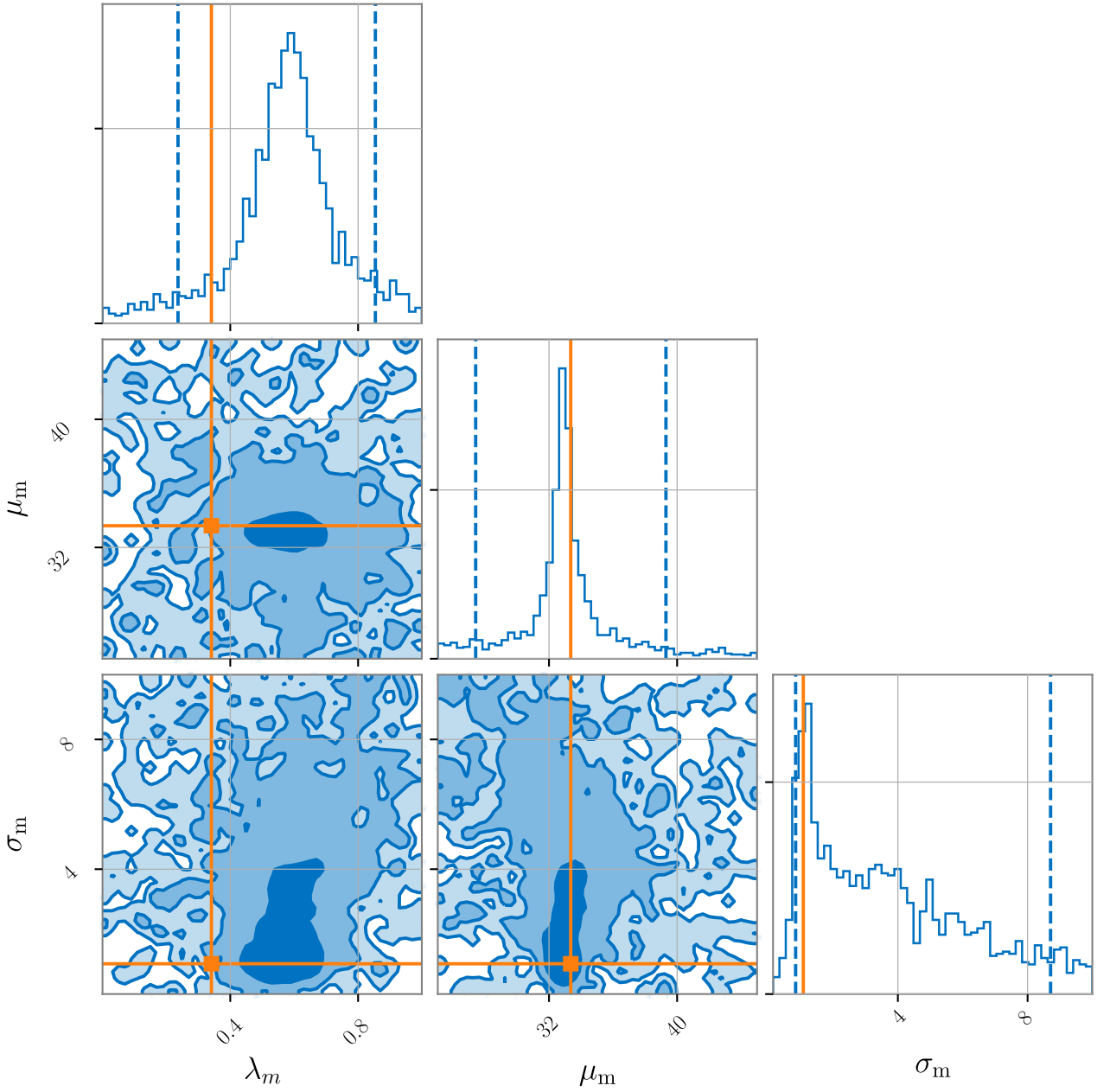
- Vitale S., Lynch R., Sturani R., Graff P., 2017, *Class. Quantum Gravity*, 34, 03LT01
- Wysocki D., Lange J., O’Shaughnessy R., 2019, *Phys. Rev. D*, 100, 043012
- You Z.-Q., Zhu X.-J., Ashton G., Thrane E., Zhu Z.-H., 2020, preprint (arXiv:2004.00036)

## APPENDIX A: POPULATION HYPER-PARAMETER ESTIMATION

The 1D and 2D PDFs for the population hyper-parameters used in this study are shown below.

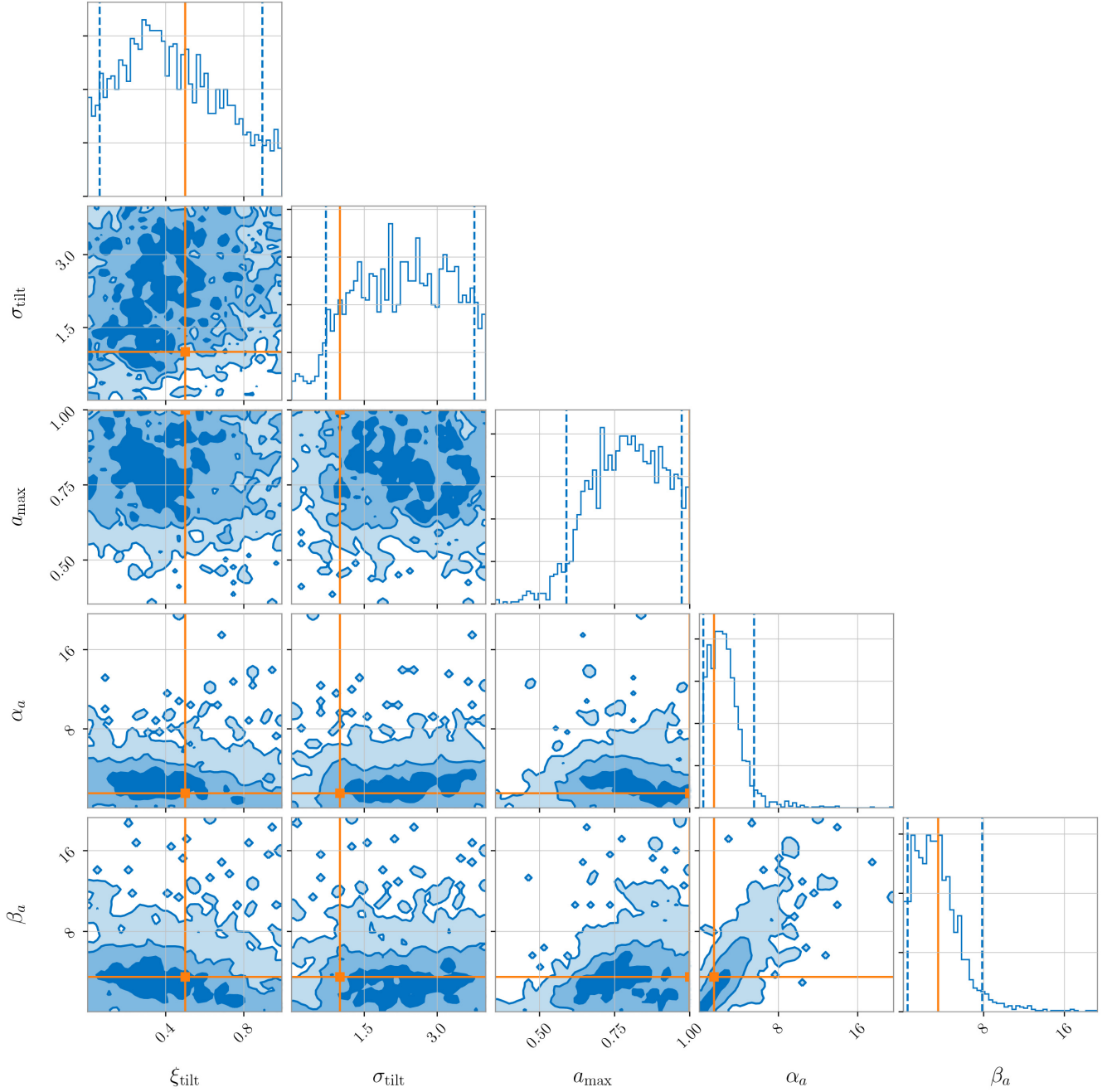


**Figure A1.** 1D and 2D (hyper-) posterior distribution. This figure showcases duty cycle  $\xi$  and hyper-parameters related to the mass-spectrum peak. From left to right; the astrophysical duty cycle  $\xi$ ; the slope of the power-law component of the primary mass distribution  $\alpha_m$ ; the slope of the mass ratio distribution  $\beta_m$ ; the minimum black hole mass  $m_{\min}$ ; and the maximum black hole mass in the power-law component  $m_{\max}$ . The dashed lines are the 90 per cent credible intervals.



**Figure A2.** 1D and 2D (hyper-) posterior distributions. This figure showcases duty cycle hyper-parameters related to shape of the BBH mass spectrum. From left to right: the fraction of black holes in the Gaussian component of the primary mass distribution  $\lambda_m$ ; the mean of the Gaussian component of the primary mass distribution  $\mu_m$ ; and the standard deviation of the Gaussian component of the primary mass distribution  $\sigma_m$ . The dashed lines are the 90 per cent credible intervals.





**Figure A3.** 1D and 2D (hyper-) posterior distributions. This figure shows hyper-parameters related to the distribution of black hole spins. From left the right: the fraction of BBHs with Gaussian distributed spin tilts  $\xi_{\text{tilt}}$ ; the standard deviation of the spin-tilt angle distribution  $\sigma_{\text{tilt}}$ ; the maximum spin magnitude  $a_{\text{max}}$ ; the spin-magnitude beta distribution slope parameter (rise)  $\alpha_a$ ; and the spin-magnitude beta distribution slope parameter (fall)  $\beta_a$ . The dashed lines are the 90 per cent credible intervals.

**APPENDIX B: POPULATION MODEL DETAILS****B1 Source-frame mass**

The conditional prior for BBH mass is:

$$\pi_m(m_1|\Lambda) = \left[ (1 - \lambda_m) A(\Lambda) m_1^{-\alpha} \Theta(m_{\max} - m_1) + \lambda_m B(\Lambda) \right] \times \exp\left(-\frac{(m_1 - \mu_m)^2}{2\sigma_m^2}\right) S(m_1|m_{\min}, \delta m),$$

$$\pi_q(q|m_1, \Lambda) = C(m_1, \Lambda) q^\beta S(m_2|m_{\min}, \delta m). \quad (\text{B1})$$

The first equation describes the prior probability of the primary mass  $m_1$  (corresponding to the heavier of the two black holes in a BBH) given the hyper-parameters  $\Lambda$ . The second equation describes the prior probability of the mass ratio  $q = m_2/m_1$  given  $m_1$  and  $\Lambda$ .

The fraction of black holes in the Gaussian component is  $\lambda_m$ . The distribution of mass ratios follows a power-law distribution with unknown spectral index  $\beta$ . Additionally, there is a smoothing parameter  $\delta m$  which enables the distribution to have a smooth turn-on at low masses.

The prior for primary mass  $\pi(m_1|\Lambda)$  is constructed from two pieces. The first term

$$(1 - \lambda_m) A(\Lambda) m_1^{-\alpha} \Theta(m_{\max} - m_1), \quad (\text{B2})$$

describes a power-law distribution with index  $\alpha \in \Lambda$ . The Heaviside step-function cuts off the distribution at  $m_{\max} \in \Lambda$ . One minus the term  $\lambda_m \in \Lambda$  is the fraction of events that are part of this power-law distribution. The term  $A(\Lambda)$  is a normalization constant. This term is motivated by the fact that the stellar mass function is power-law distributed as well as evidence of a cut-off in the black hole mass spectrum (Fishbach, Holz & Farr 2017; Talbot & Thrane 2017; Abbott et al. 2018a).

The second term in  $\pi(m_1|\Lambda)$

$$\lambda_m B(\Lambda) \exp\left(-\frac{(m_1 - \mu_m)^2}{2\sigma_m^2}\right), \quad (\text{B3})$$

corresponds to a Gaussian distribution with mean  $\mu_m \in \Lambda$  and width  $\sigma_m \in \Lambda$ . The fraction of events that are part of the Gaussian distribution is given by  $\lambda_m$ . The  $B(\Lambda)$  term is a normalization constant. This term is motivated by the possibility of a bump in the black hole mass spectrum from pulsational pair instability supernovae (Abbott et al. 2018a; Talbot & Thrane 2018; Marchant et al. 2019).

To the far right of the expression for  $\pi(m_1|\Lambda)$  is a third term

$$S(m, m_{\min}, \delta m) = (\exp f(m - m_{\min}, \delta m) + 1)^{-1}$$

$$f(m, \delta m) = \frac{\delta m}{m} - \frac{\delta m}{m - \delta m}. \quad (\text{B4})$$

The  $m_{\min}$  parameter enforces a minimum black hole mass and  $\delta m$  is the mass range over which the black hole mass spectrum falls to zero. This term is motivated by the fact that there is likely a minimum black hole mass, at least for black holes made through stellar collapse (Talbot & Thrane 2018).

The conditional prior for mass ratio is described by a power law with index  $\beta \in \Lambda$ . The smoothing function  $S$  applies a low-mass

cut-off in the secondary mass  $m_2$ , again using minimum mass  $m_{\min}$  and  $\delta m$  for the mass range over which the mass spectrum falls to zero. The variable  $C(m_1, \Lambda)$  is a normalization constant.

**B2 Lab-frame mass**

The BBH lab-frame mass is a function of redshift because

$$m_l = (1 + z)m_s, \quad (\text{B5})$$

where  $m_s$  is the source-frame mass and  $m_l$  is the lab-frame mass. When considering events at cosmological distances, the prior distributions for lab-frame masses become covariant with luminosity distance  $D_L$  due to cosmological redshift. In the source frame, the distributions of black hole mass and redshift are separable so that

$$\pi(m_s, z) = \pi_m(m_s) \pi_z(z) \quad (\text{B6})$$

Whatever form the distributions we choose for  $\pi_z(z)$  and  $\pi_m(m_s)$ , they imply some prior for the lab-frame mass:

$$\pi(z, m_l) = \pi(z, m_s(m_l)) \left| \frac{dm_s}{dm_l} \right|$$

$$= (1 + z)^{-1} \pi(z, m_l/(1 + z)). \quad (\text{B7})$$

**B3 Spin**

The distribution of spin magnitudes ( $a_1, a_2$ ) are assumed to each follow a beta distribution described by three parameters ( $\alpha_a, \beta_a, a_{\max}$ )  $\in \Lambda$ . By treating  $a_{\max}$  as a free parameter, our model is a generalization of the prescription from Wysocki et al. (2019). The conditional prior for spin magnitude is

$$\pi_a(a|\alpha_a, \beta_a, a_{\max}) = \frac{a^{(\alpha_a-1)}(a_{\max}-a)^{(\beta_a-1)}}{a_{\max}^{(\alpha_a+\beta_a-1)} B(\alpha_a, \beta_a)}. \quad (\text{B8})$$

Here,  $B(\alpha_a, \beta_a)$  is the Beta function.

We characterize the black hole spin orientation in terms of the cosine of the polar angle between the orbital angular momentum and the black hole spin  $z_{1,2} \equiv \cos(\theta_{1,2})$  where  $\theta_{1,2}$  is the polar angle. We ignore the azimuthal angle, which has a comparatively small effect on the gravitational waveform. We assume that the distribution of spin orientations is a mixture of an isotropic component and a preferentially aligned component modelled as a truncated half-Gaussian with unknown width  $\sigma_{\text{tilt}}$  and which peaks at  $t_1 = t_2 = 1$ .

$$\pi(z_1, z_2|\sigma_{\text{tilt}}, \lambda_{\text{tilt}}) = \frac{(1 - \lambda_{\text{tilt}})}{4}$$

$$+ \frac{\lambda_{\text{tilt}}}{2\pi} \prod_{i \in \{1,2\}} \frac{e^{-(1-z_i)^2/(2\sigma_{\text{tilt}}^2)}}{\sigma_{\text{tilt}} \text{erf}(\sqrt{2}/\sigma_{\text{tilt}})}. \quad (\text{B9})$$

The isotropic distribution is a model for mergers in dense stellar environments such as globular clusters, where spin orientations are expected to be isotropically oriented. The aligned distribution models binaries formed in the field. The fraction of binaries in the preferentially aligned component is  $\xi_\chi$ . We assume that both component spins are independently drawn from the same distribution.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.